Detection and recognition of pure tones in noise

David M. Green, Daniel L. Weber, and Joseph E. Duncan

Laboratory of Psychophysics, Harvard University, Cambridge, Massachusetts 02138

(Received 11 April 1977; revised 10 June 1977)

We examine the predictions of a new theorem relating signal identification (specifying a signal as a particular member of a set of potential signals) to signal detection (discriminating the presence of a signal). The theorem, derived in the context of signal-detection theory, requires that the signals be equally detectable and orthogonal. Our sinusoidal signals are partially masked by noise and their intensities adjusted to produce equal-signal detectability; we do not examine this assumption of the theorem. The theorem generally provides a reasonably accurate description of recognition performance for two-signal and four-signal conditions and is equally accurate for both the Yes-No and category-rating procedures. In a preliminary investigation of the orthogonality assumption, we varied the frequency separation between two signals. When the frequency separation between two signals is small (20 Hz near 1 kHz), the theorem fails to provide a good description of performance.

PACS numbers: 43.66.Ba, 43.66.Dc, 43.66.Fe, 43.66.Lj

INTRODUCTION

Detection is usually described as the ability to discriminate the addition of some signal to either a quiet or noisy background. Recognition, identification, or localization implies the additional ability to select the detected signal from a set of many. A theorem recently established by Starr, Metz, Lusted, and Goodenough (1975) predicts recognition performance from detection performance. Their visual task requires observers to state whether a tumor is present in an x ray (detection), and if it is, to identify which quadrant of the x ray it is in (localization or recognition). Their theorem, developed in the framework of signal detection theory, is formally applicable to auditory detection tasks. Our three experiments seek to determine conditions under which the theorem is valid.

All signals are pulsed sinusoids, one of several different frequencies, partially masked by noise. The observer gives two responses, first a detection response—whether a signal is present or not—and second, a recognition response—which of several possible signals it might be. In three different experiments we examine the effect of experimental method (Yes-No versus rating procedures), number of signals, and frequency separation between signals. A strong assumption of the theorem, the orthogonality of the signals, is tested by this last experiment. The second strong assumption in the theorem, equal signal detectability, is not experimentally examined; we tried to make all signals approximately equal in detectability.

I. STATEMENT OF THEOREM

To understand the recognition theorem, we need to explain the experimental conditions in more detail. Consider the following situation: A single observation interval contains either the noise alone, or one of m possible signals added to the noise. The observer first tries to identify the signal by giving one of m recognition responses. First we analyze the detection response in terms of an ROC curve—a plot of hit versus false alarm probability. Although constructed by the usual method, it is not a conventional ROC curve since all m potential signals are treated alike; for this reason we will refer to it as a 1-of-m ROC curve. As in the conventional ROC curve, each point on the 1-of-m ROC is associated with a different detection criterion, \( \lambda \). Thus we write the hit rate as \( P(Y|s, \lambda) \) and the false-alarm rate as \( P(Y|n, \lambda) \).

Now consider the ratio

\[
R(\lambda) = \frac{1 - P(Y|s, \lambda)}{1 - P(Y|n, \lambda)}
\]

generated by a particular decision criterion, \( \lambda \). The numerator of \( R(\lambda) \) is the vertical distance between the given point and the point (1, 1). The denominator is the horizontal distance between the same two points. The ratio, then, is simply the slope of a line drawn through the point on the 1-of-m ROC curve and the point (1, 1) [Fig. 1(a)]. This observation makes it easy to see that as the criterion \( \lambda \) is varied from a very strict to a very lenient one, the ratio \( R(\lambda) \) will start with a value of 1 and decrease monotonically for most ROC curves [Fig. 1(b)].

The recognition theorem asserts that the expected probability of a correct recognition following an affirmative detection response for the set of \( m \) signals, \( P_m(c \cdot Y|\lambda) \), is

\[
P_m(c \cdot Y|\lambda) = P(Y|s, \lambda) \frac{1 - P(Y|s, \lambda)}{1 - P(Y|n, \lambda)} \int_0^{P(Y|n, \lambda)} dP(Y|n, \lambda)
\]

for a given detection criterion, \( \lambda \).

The theorem predicts that if detection performance is at chance, then recognition performance is also at chance. For \( P(Y|s, \lambda) = P(Y|n, \lambda) \), i.e., chance detection performance, \( R(\lambda) = 1 \), and the integral reduces to \( P(Y|n, \lambda) \), which is equal to \( P(Y|s, \lambda) \), and thus the entire equation reduces to \((1/m)P(Y|s, \lambda)\)—one guesses among the \( m \) signals with equal probability. If one's detection performance is very good, recognition perfor-
II. PROCEDURE

Two observers sat in an IAC (Model 402A) sound-treated chamber and listened through Superex head-phones. There were four undergraduate observers. Observers JU and LT were tested as one pair, observ-ers GR and KL were tested as another pair.

The spectrum level of the continuous noise masker was 46 dB. Both the noise and the signals were pre-sented binaurally. The probability of signal and of no-signal trials was one half. For signal trials, signal frequency was chosen with equal probability from the signal set. The signal-to-noise ratio ($E/N_0$) for each frequency was adjusted in an attempt to equate the detectability at each signal frequency. The formula $10 \log E/N_0 = 2(f/f_0) + 8 \text{dB}$, where $f$ is the frequency of the signal in hertz and $f_0=1000$ Hz, was used as a first approximation (Green, McKey, and Licklider, 1959). Slight adjustments ($\pm 1 \text{ dB}$) were then made to better equate the detectability at each frequency for the particular pair of observers. The signal levels at all frequencies were about 1.5 dB lower for JU and LT than for GR and KL.

About 3–5 days of practice preceded any data collection. The initial trials were run at very large signal-to-noise ratios so that the observers could become familiar with the task and correctly identify various frequencies. In the experiment with four signal frequencies all observers could obtain 100% recognition scores if the noise were lowered to a spectrum level of 6 dB. Typically, six blocks of 100 trials were completed in each daily 2-h session. The observers saw a summary of their performance during the 5–10 min rest between blocks.

A PDP-15 computer controlled all experimental operations; each observer used an individual response box with a 16-key calculator keyboard and LED (light-emitting diode) display interfaced to the computer. Each trial began with a 1-sec warning light presented on the observer’s LED display. The termination of the warning light was followed by a 100-msec observation interval (the signal duration), during which time a signal might be presented. The observer then gave the detection response, either Y–N (1 or 6) or rating (1, 2, 3,..., 6). The response could be modified by pressing a control key and reentering another response. The observer finally pressed another control key which caused the re- sponse to be read by the computer and in doing so cleared the display. The observer then entered a rec-ognition response in essentially the same manner as the detection response. After the recognition response was recorded, there was a 500-msec feedback interval (0 = no signal; digit 1 to $m$ = signal 1 to $m$ presented). After a 500-msec pause, a new trial began.

III. RESULTS AND DISCUSSION

A. Yes–No data with four potential frequencies—

Experiment 1

The possible signal frequencies were 500, 1100, 1900, and 2700 Hz. On each trial the observer first gave a detection response (Yes or No) and then the recognition response. A recognition response was required even if the detection response was negative. We instructed one pair of observers to adopt three different criteria in different blocks of trials; the other observers we instructed to adopt six different criteria, varying from very conservative to very lenient. False alarms ranged from 0.01 to 0.78 for the different conditions. Figure 2 shows the data from this experiment for four observers. The solid, square points are the observed proportions of hits and false alarms for the detection task. Each point is
TABLE I. Experiment 1. Yes–No procedure.

<table>
<thead>
<tr>
<th>False alarm probability</th>
<th>Probability (correct recognition</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–0.10</td>
<td>0.454</td>
<td>530</td>
</tr>
<tr>
<td>0.10–0.29</td>
<td>0.376</td>
<td>246</td>
</tr>
<tr>
<td>0.30–0.50</td>
<td>0.337</td>
<td>181</td>
</tr>
<tr>
<td>0.60–1.00</td>
<td>0.319</td>
<td>97</td>
</tr>
</tbody>
</table>

Based on about 300 observations: 150 signal-plus-noise and 150 noise-alone trials. The observed proportion of correct recognition responses is plotted as a solid circle at the false alarm probability corresponding to that detection criterion. This proportion is the average of the diagonal entries of the m by m stimulus-response matrix obtained at the given criterion. The open circular points are the proportion of correct recognition responses predicted on the basis of the detection responses.

As can be seen, the obtained data falls systematically below the predicted point for observers JU, GR, and KL, and nearly on the predicted value for LT. The average discrepancy is about 8% at the most lenient false alarm rate and smaller at the more strict criteria.

As mentioned above, the observer gave a recognition response even when the detection response was negative. Table I gives the percentage of correct recognition responses given a "No" detection response. As one might expect, the recognition percentage for the very strict criteria is considerably above the chance level, 0.45 rather than 0.25. As the criteria are made more lenient, this probability decreases toward the chance level.

B. Rating data with four potential frequencies—Experiment 2

The same four observers listened to the same four signal frequencies but gave a six-point rating as their detection response prior to their recognition response. Figure 3 shows the results of this experiment. Once again the 1-of-m ROC curve (solid squares) yields an estimate of R(λ) and we thus predict recognition accuracy (open circle) for each point. The obtained recognition score is indicated by the solid circle. Again, the obtained scores for three of the four observers are systematically lower than the predicted values. The discrepancy is about the same size as in Experiment 1. 1-of-m ROC, o-predicted, and •-obtained.

For each observer we can compute the probability of a correct recognition response given each criterion level. These conditional probabilities are given in Table II. For each observer there is a monotonic increasing relation between the detection category and the probability of a correct recognition. For the most lenient criterion (almost certain there was no signal) the recognition accuracy is nearly at the chance level, but all observers are slightly greater than 0.25.

The theorem assumes equal detectability, and during the training period, we attempted to find signal levels which produced this condition. To check this, we can tabulate the probability of correct responses given each signal. Table III contains the data from this analysis for the four observers. All observers seem somewhat better at detecting the extreme signals, 500 and 2700 Hz, than the middle signals, 1100 and 1900 Hz. There are substantial response biases that can be observed by computing the percentage of time each response occurred given noise alone. GR and KL tended to give middle responses, 1100 and 1900 Hz, more (55%) than extreme responses, 500 and 2700 Hz (45%). JU had the reverse pattern and LT had essentially no biases. Interpreting these differences is difficult, since the observers were given feedback after each block of 100 trials. If the observers felt one frequency was difficult
to hear, they might respond with that frequency whenever they were uncertain, and hence increase the percentage of correct responses at that frequency.

C. Rating data with two potential signals—Experiment 3

The rating experiment with four signals is a difficult task. We decided to simplify the task by using only two frequencies. In addition we varied the difference in frequency between the two signals and thereby explored the question of whether violation of the orthogonality assumption influenced the discrepancy between the predicted and the obtained recognition scores. The rating method was used throughout since it seemed to work as well as the Yes-No task and was more efficient. The first two frequencies tested were 500 and 2700 Hz. Figure 4 shows the data obtained from about 1000 trials. The predicted and obtained scores are in nearly perfect agreement for JU and LT; for GR and KL the obtained recognition scores are somewhat lower than we would predict given the detection performance. Table IV presents the recognition score for each rating category. For three of the observers the recognition score is above chance even at the lowest rating. In general, the recognition accuracy increases monotonically with detection category, as one might expect. Note GR and KL, who obtain lower recognition scores than we would predict, have very high recognition scores for the lowest detection category. Observer KL, for example, achieves a recognition probability of 0.83 ($N = 23$) on those trials when he was almost certain the signal had not been presented. We are inclined to believe GR and KL did not use the ratings in the detection part of the task as most other observers do.

Next the two signals were moved closer together in frequency, 940 and 1060 Hz, to be slightly less than a critical band apart according to some estimates (Zwicker, Flottrup, and Stevens, 1957), and slightly larger than a critical band according to others (Fletcher, 1940). Figure 5 shows the 1-of-$m$ ROC curve and predicted and obtained recognition scores; the predictions are now somewhat lower than the obtained results for GR and KL. Table V shows how recognition accuracy varied with the detection category. There is still a nearly monotonic relation between recognition accuracy and detection category.

Finally, we brought the frequencies very close together, 990 and 1010 Hz, well within any estimates of critical band. The results are given in Fig. 6 and Table VI. As can be seen, the obtained recognition scores decrease by about 20% for all observers, falling below the predicted scores by sizeable amounts for observers JU and LT.

IV. DISCUSSION

Conceptually we may think of the signals being discriminated by different processes or detectors. If the outputs of these detectors are independent, then the probability of any combined outputs can be calculated by multiplying the probabilities of each separate output. When the detectors are independent one can describe the signals as being orthogonal, although this usage is not

<table>
<thead>
<tr>
<th>Detection category</th>
<th>JU</th>
<th>LT</th>
<th>GR</th>
<th>KL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.55</td>
<td>0.45</td>
<td>0.57</td>
<td>0.64</td>
</tr>
<tr>
<td>2</td>
<td>0.52</td>
<td>0.61</td>
<td>0.68</td>
<td>0.65</td>
</tr>
<tr>
<td>3</td>
<td>0.55</td>
<td>0.65</td>
<td>0.80</td>
<td>0.79</td>
</tr>
<tr>
<td>4</td>
<td>0.59</td>
<td>0.76</td>
<td>0.93</td>
<td>0.81</td>
</tr>
<tr>
<td>5</td>
<td>0.68</td>
<td>0.93</td>
<td>0.97</td>
<td>0.95</td>
</tr>
<tr>
<td>6</td>
<td>0.63</td>
<td>0.99</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

FIG. 4. Rating ROC for wide separation of two signal frequencies. Signal frequencies are 500 and 2700 Hz. •-1-of-$m$ ROC, o-predicted, and e-obtained.
technically precise. Sometimes signals can be represented in a vector space in which case orthogonal signals are those whose inner products are zero. If we knew more about the detection process we might be able to compute a quantity similar to the inner product at the outputs of the audition detectors, and directly evaluate the relation between any two signals. At this time no one is able to give an adequate representation of the signal. Nonetheless, we can provide a precise test of the independence (within the framework of the theorem) of the several detection processes and thus contribute to the understanding of when signals are best represented as orthogonal or nonorthogonal.

Ignoring the last experimental condition, where the independence assumption is surely wrong, the theorem generally provides reasonably good predictions of recognition performance. Consistent individual differences exist among the observers but, if we average over observers, the theorem’s predictions would be very close to the obtained results. Lindner (1968) examined the relation between detection and recognition performance using 500 and 1100 Hz sinusoidal signals and four different response criteria. The stimulus conditions are similar to those we employed in Experiment 3; he, however, used a Yes-No design rather than a rating procedure. The recognition theorem predicts values very close to his results, the average error (over all observers and conditions) is about 1%.

The average error for our first experiment was +0.02 (an average absolute difference of 0.045). The average error for the second experiment was +0.03 (an average absolute difference of 0.04). Although simple tests cannot reject the hypothesis that the error is zero (the null hypothesis), perhaps a more sensitive experiment, with more observations per condition, would do so. Of more interest is to compare these predictions with those generated by alternative theories, but as yet, no such predictions are available.

There is a systematic decrease in recognition performance, relative to the predicted value, as the frequency separation between signals decreases (Experiment 3). Performance falls markedly below the predicted value when the separation is small. This failure is easily interpreted as a violation of the theorem’s orthogonality assumption. The 20-Hz separation is much less than any estimate of the critical band, and the signals in this condition clearly do not have independent representations in the auditory system. The average error in prediction is about 0.11 for this condition and hence the procedure does provide some measure of orthogonality. This procedure can be used to test the orthogonality of two signals, a clear advantage over the analysis of confusion matrices or ratings of similarity, such as those used in multidimensional scaling procedures, where at least three signals must be compared.

The accuracy of the predictions appears to be equally good for both two- and four-frequency conditions and to be relatively unaffected by the mode of detection response used (Yes-No versus rating) even though there are differences in the size of the signal set in the first comparison and differences in the size of the response set in the second. These results are all consistent with the theorem and the fact that the rating procedure works as well as the Yes-No procedure has practical significance, since it is considerably more efficient.

In summary, we have examined a new theorem relating recognition performance to detection performance using sinusoidal signals as elements of the recognition set. The theorem provides accurate predictions except for a small frequency separation. Future theoretical exploration will be needed before applying this approach to signals of unequal detectability. Another avenue for future research involves the application of the theorem to more complex signals.

ACKNOWLEDGMENTS

Application of the theorem proved by Starr, Metz, Lusted, and Goodenough to recognition and classification tasks was initiated by the senior author and John A. Swets, at Bolt Beraneck and Newman Inc., in research sponsored by the Engineering Psychology Programs of the Office of Naval Research. The research in this paper was supported in part by the National Institutes of Health, Public Health Service, U.S. Department of Health, Education, and Welfare. We also wish to thank Judith Ullman, who collected part of the data, and James Egan, William Kelly, Walt Jesteadt, Charles Metz, David Noreen, Neal Viemeister, Charles Watson, and Craig Wier, all of whom commented on earlier drafts of this paper.

TABLE VI. Experiment 3. Two signals, small separation probability of a correct recognition for each rating category.

<table>
<thead>
<tr>
<th>Detection rating</th>
<th>JU</th>
<th>LT</th>
<th>GR</th>
<th>KL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.65</td>
<td>0.23</td>
<td>0.53</td>
<td>0.47</td>
</tr>
<tr>
<td>2</td>
<td>0.54</td>
<td>0.40</td>
<td>0.57</td>
<td>0.56</td>
</tr>
<tr>
<td>3</td>
<td>0.50</td>
<td>0.51</td>
<td>0.50</td>
<td>0.57</td>
</tr>
<tr>
<td>4</td>
<td>0.54</td>
<td>0.65</td>
<td>0.55</td>
<td>0.66</td>
</tr>
<tr>
<td>5</td>
<td>0.42</td>
<td>0.62</td>
<td>0.58</td>
<td>0.65</td>
</tr>
<tr>
<td>6</td>
<td>0.60</td>
<td>0.78</td>
<td>0.75</td>
<td>0.64</td>
</tr>
</tbody>
</table>

FIG. 6. Rating ROC for small separation between two signal frequencies. Signal frequencies are 990 and 1010 Hz. -1-of- m ROC, o-predicted, and o-obtained.
APPENDIX

We now derive the relation between the recognition score and the hit and false alarm probabilities of the 1-of-m ROC curve. We use capital letters to denote the coordinates of the 1-of-m ROC curve. We can consider the 1-of-m detection decision as composed of m tests carried out on m independent, elementary detectors. The hit and false alarm probabilities of these more elementary detectors we denote with small letters. Hence $P(Y | n, \lambda)$ is the false-alarm probability for the 1-of-m ROC curve given a criterion $\lambda$, and $p(Y | n, \lambda)$ is the false alarm probability associated with one of the m elementary tests. We can also represent this latter probability as the area above some noise-alone distribution,

$$p(Y | n, \lambda) = \int_{-\infty}^{\infty} f(x | n) dx .$$

If the detectors are independent then the probability of no false alarm, given noise alone, is simply the product of the probability of no false alarm for each of the m detectors.

$$1 - P(Y | n, \lambda) = [1 - p(Y | n, \lambda)]^m = \left[ \int_{-\infty}^{\infty} f(x | n) dx \right]^m. \quad (A1)$$

The formula for the hit probability is similar,

$$1 - P(Y | s, \lambda) = \left[ 1 - p(Y | s, \lambda) \right] \left[ 1 - p(Y | n, \lambda) \right]^{m-1} = \left[ \int_{-\infty}^{\infty} f(x | s) dx \right] \left[ \int_{-\infty}^{\infty} f(x | n) dx \right]^{m-1}. \quad (A2)$$

Using this last expression we obtain

$$\int_{-\infty}^{\infty} f(x | s) dx = [1 - P(Y | s, \lambda)] \int_{-\infty}^{\infty} f(x | n) dx. \quad (A3)$$

Differentiating both sides with respect to $\lambda$, the criterion, we find

$$f(\lambda | s) = \frac{-dP(Y | s, \lambda)}{d\lambda} \left[ \int_{-\infty}^{\infty} f(x | n) dx \right]^{m-1} - \left[ 1 - P(Y | s, \lambda) \right] (m-1) \left[ \int_{-\infty}^{\infty} f(x | n) dx \right]^{m-2} f(\lambda | n). \quad (A4)$$

We assume that the recognition response is given to the detector with the maximum output and that a detection response is made whenever the output of at least one detector exceeds the detection criterion. To correctly recognize a signal after saying "Yes" thus requires that the following two conditions be satisfied.

First, the observation associated with the signal detector [X in Eq. (A5)] exceed the m-1 observations associated with all the noise-alone detectors. Second, the observation must exceed $\lambda$, so that an affirmative detection response occurs. The probability of correctly saying "Yes" and recognizing the signal, denoted $P_m(c \cdot Y | \lambda)$, may then be written, using the orthogonality assumption, as

$$P_m(c \cdot Y | \lambda) = \int_{-\infty}^{\infty} f(\lambda | s) \left[ \int_{-\infty}^{\infty} f(x | n) dx \right]^{m-1} d\lambda . \quad (A5)$$

Substituting $f(\lambda | s)$ from Eq. (A4) in Eq. (A5) and simplifying yields

$$P_m(c \cdot Y | \lambda) = \int_{-\infty}^{\infty} \left[ -dP(Y | s, \lambda) \left[ \int_{-\infty}^{\infty} f(x | n) dx \right]^{m-1} - \left[ 1 - P(Y | s, \lambda) \right] (m-1) \left[ \int_{-\infty}^{\infty} f(x | n) dx \right]^{m-2} f(\lambda | n) \right] d\lambda . \quad (A6)$$

Changing the sign of the first term and reversing the limits of integration lets us see that it is simply $P(Y | s, \lambda)$. Using Eq. (A1) and differentiating, we have

$$-dP(Y | n, \lambda) = m \left[ \int_{-\infty}^{\infty} f(x | n) dx \right]^{m-1} f(\lambda | n) d\lambda. \quad (A7)$$

Making the appropriate changes in the limits of integration and substituting $f(\lambda | n)$ in the right-most term of Eq. (A6) we have

$$P_m(c \cdot Y | \lambda) = P(Y | s, \lambda) - \frac{m-1}{m} \int_{-\infty}^{\infty} \left[ 1 - P(Y | s, \lambda) \right] \frac{1}{1 - P(Y | n, \lambda)} dP(Y | n, \lambda). \quad (A8)$$

This is the recognition theorem presented by Starr et al. We will now derive the area theorem as a special case of the recognition theorem. In the limit as $\lambda \to -\infty$, i.e., when we always say yes, Eq. (A8) becomes

$$P_m(c) = 1 - \frac{m-1}{m} \int_{-\infty}^{\infty} \left[ 1 - P(Y | s, \lambda) \right] \frac{1}{1 - P(Y | n, \lambda)} dP(Y | n, \lambda). \quad (A9)$$

Note that

$$1 - P(Y | s, \lambda) = \frac{[1 - p(Y | s)] [1 - p(Y | n)]^{m-1}}{1 - P(Y | n, \lambda)} = \frac{1 - p(Y | s, \lambda)}{1 - p(Y | n, \lambda)} .$$

Using Eq. (A7), as well as this identity, we have

$$P_m(c) = 1 - \frac{m-1}{m} \int_{-\infty}^{\infty} [1 - p(Y | s)] [1 - p(Y | n)]^{m-2} dP(Y | n).$$

Using integration by parts where $u = 1 - p(Y | s)$ and $v = [1/(m-1)][1 - p(Y | n)]^{m-1}$ we have

$$P_m(c) = 1 - \frac{m-1}{m-1} \left[ 1 - p(Y | s) \right] [1 - p(Y | n)]^{m-1} \int_{0}^{1} dP(Y | s).$$

When evaluated, this yields,
\[ P_m(c) = \int_0^1 [1 - p(Y|n)]^{m-1} dp(Y|s), \]

which is the generalized form of the area theorem (see Green and Swets, 1974, p. 47). If \( m = 2 \), this is the familiar result that the area under the Yes-No ROC curve is the percent correct in the two-alternative forced-choice task.

\( \text{The } j \text{ points on the 1-of-} m \text{ ROC curve yield } j \text{ discrete estimates of } R(\lambda) \text{ for } j \text{ values of } P(Y|n, \lambda). \text{ One must use these discrete estimates to approximate the integral of } R(\lambda). \text{ This involves the piecewise approximation of the integral between successive } P_i(Y|n, \lambda). \text{ Since } R(\lambda) \text{ is, in general, a monotonically decreasing function, } R_i(\lambda)[P_i(Y|n, \lambda) - P_{i+1}(Y|n, \lambda)] \text{ underestimates the piece of the integral between } P_i(Y|n, \lambda) \text{ and } P_{i+1}(Y|n, \lambda). \text{ Using } R_{i+1}(\lambda) \text{ results in overestimation. We used the mean of successive values of } R(\lambda), \frac{1}{2}[R_i(\lambda) + R_{i+1}(\lambda)], \text{ to evaluate the piecewise integral at } P_i(Y|n, \lambda). \text{ The approximation of the entire integral to } P(Y|n, \lambda) = 0 \text{ was used for evaluating the first piecewise integral, i.e., for } i = 1. \text{ The limit of } R(\lambda) = 1 \text{ for } P(Y|n, \lambda) = 0 \text{ was used for evaluating the first piecewise integral, i.e., for } i = 1. \text{ Green, D. M., McKey, M. J., and Licklider, J. C. R. (1959). "Detection of a pulsed sinusoid in noise as a function of frequency," J. Acoust. Soc. Am. 31, 1446–1452.}


